DOCUMENT RESUME

ED 423 272 TM 029 077

AUTHOR Centra, John A.; Gaubatz, Noreen B.
TITLE Is There Gender Bias in Student Evaluation of Teaching?
PUB DATE 1998-04-00
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Bias; *College Students; Females; Higher Education; Males; *Sex Differences; *Student Evaluation of Teacher Performance; *Teacher Effectiveness; Teaching Methods

ABSTRACT

This study examined gender differences in student evaluation of teaching through two analyses. In the first, female and male student ratings in the same classes were compared for female and male instructors. In the second analysis, ratings by all male students are examined for how they differed for male and female instructors. Data came from 741 college classes each of which had an enrollment of at least 10 female students and 10 male students from 21 colleges and universities. The student evaluation form was the Student Instructional Report II from the Educational Testing Service. Multivariate analysis of variance was used to investigate the mean differences on the dependent variables. In this study, in contrast to past studies, female students gave higher ratings to female instructors than to male instructors on three of eight scales for all disciplines combined, while male students gave male instructors higher ratings only on one scale, course organization and planning. Male and female students did not differ in their rankings of male teachers. For the total sample of classes, when more favorable ratings were given, they were largely by female students to female instructors. Overall, results support the conclusion that gender differences among instructors are related more to their gender-related approaches to teaching than to their overall effectiveness. (Contains 4 tables and 15 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 423 272

Is There Gender Bias in Student Evaluation of Teaching?

John A. Centra
Noreen B. Gaubatz

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
John Centra
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
2B

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

This paper is prepared for the:
Annual Meeting of the American Educational Research Association in San Diego, CA
April 1998

TM029077

IS THERE GENDER BIAS IN STUDENT EVALUATIONS OF TEACHING?[1]

John A. Centra and Noreen B. Gaubatz

Given the widespread use of student evaluations of teaching for tenure and promotion decisions, it is important to be aware of possible bias in the evaluations. Bias is commonly defined in this instance as a teacher or course characteristic that affects teacher evaluations, either positively or negatively, but is unrelated to criteria of good teaching, such as increased student learning. Class size, for example, affects evaluations in that classes with under 15 students get higher evaluations. But if students learn more than they do in larger classes, perhaps because small classes allow for more personal attention, then class size is not truly biasing the evaluations; rather the evaluations are reflecting increased learning.

The gender of the student, in particularly how it interacts with the gender of the teacher, is a possible source of bias in student evaluations that has not been fully investigated. Do male students tend to rate women teachers lower than men teachers because of a sex bias, especially in fields that are male dominated such as the natural sciences. Do female students judge women teachers to be more effective than men, even though they may learn no more from them. These are important questions that directly affect the validity of the evaluations when used for personnel decisions. Ideally student evaluations should be related to what they learn from a teacher and not gender or other personal characteristics of the teacher (e.g. age, ethnicity).

Studies that have investigated gender bias have thus far produced conflicting results. Some studies have found no (or extremely small) differences between the

---

[1]Presented at the American Educational Research Association meeting in San Diego, April, 1998.

evaluation of female and male instructors on the basis of student gender alone (Basow and Distenfeld, 1985; Basow and Howe, 1987; Bennett, 1982; Elmore and LaPoint, 1974; Harris, 1975; Kaschak, 1981) Other studies reported gender bias, with male students rating female instructors lower than male instructors (Basow and Silberg, 1987; Etaugh and Riley, 1983; Kaschak, 1978; Lombardo and Tocci, 1979; Paludi and Bauer, 1983).

Two major studies conducted in actual classrooms did not report gender bias in the evaluation process. Bennett's (1982) study used a course evaluation questionnaire that included teaching performance ratings, perceptual orientation scales, and indicators of the degree and context of student-instructor interaction. Her data included the evaluations of 11 female and 28 male instructors by 253 students enrolled in nonscience introductory courses at a liberal arts college. Female and male students did not differentiate between faculty members of different gender. Although there was no evidence of direct bias in formal student evaluations of instructors, there was evidence of gender related differences in regard to student-instructor interactions and instructor warmth, support, and accessibility. Elmore and LaPoint (1974) found no interaction between faculty gender and student gender in their analysis of 38 pairs of courses (paired on the basis of course number and gender of instructor) evaluated by 1,259 students. The courses were selected from a variety of departments and colleges within one research university. No attempts were made, however, to compare results from various disciplines or results for students in the same classes.

Research that reported gender bias include Basow and Silburg (1987) study of 16 pairs of instructors (paired on rank, discipline and years of experience), in which male students rated female instructors less favorably than male instructors. Similar results by Kaschak (1978) and Lombando and Tucci (1979) were found in a simulated rather than actual classroom setting. In these studies female students saw no

difference in the effectiveness of male and female teachers.

Twenty-eight studies that addressed gender bias in teacher evaluations were used in a meta-analysis by Feldman (1993). In many of these studies female and male instructors were not evaluated very differently from a global or overall perspective, but there were differences between the genders on specific aspects of teaching, such as gender stereotypic behaviors and teaching methodology.

The question of gender bias in the evaluations of instructors may not have been fully resolved because past studied relied on small n's and failed to control for significant variables, such as academic discipline. The present study addressed many of the limitations found in the current research. It used actual student ratings of classroom instructors, which reduced concerns that simulated designs lack validity (Basow and Howe, 1987; Harris, 1975; Kaschak, 1978, 1981). A large number of different types of institutions (two-year and four-year colleges and universities) were included, while many of the past studies were limited to a single institution (Basow and Silberg, 1987; Bennett, 1982; Elmore and LaPointe, 1974). A variety of academic disciplines rather than one or a relatively small number of disciplines were analyzed. Finally, unlike previous studies, the class rather than the individual student was the unit of analysis, thereby increasing reliability and minimizing the effects of extraneous students variables.

Method

This study examined gender differences through two analyses. In the first, female and male student ratings in the same classes were compared for female instructors and for male instructors. This analysis more directly addressed the purpose of this study because it compared ratings by students of the same instructors. Figure 1 illustrates this analysis.

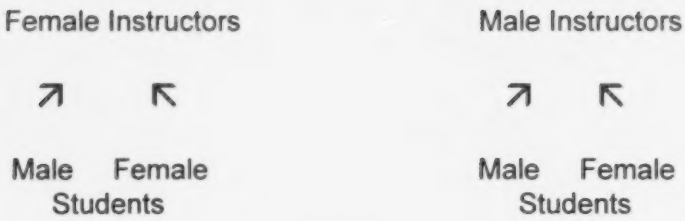

Figure 1

In the second analysis the ratings by all male students are examined for how they differed for male and female instructors. The ratings by all female students are compared in the same way. Figure 2 illustrates this analysis. Because the male or female students have evaluated different instructors, this analysis focuses more on student gender effects and their interactions with instructor gender effects.

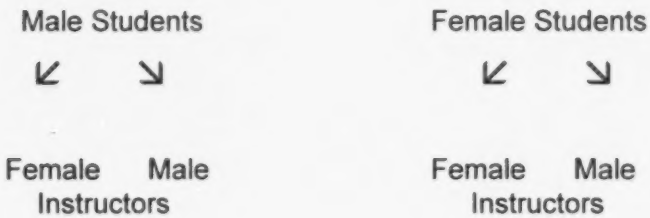

Figure 2

Subjects

The data in this study included 741 classes, each of which had an enrollment of at least 10 female students and 10 male students. The minimum number of students was

based on Centra's (1998) research indicating that scores based on 10 or more students provide a sufficient level of reliability for research purposes. Specifically, the intraclass reliabilities were in the high 50's for individual items; scale means were used in this study and they would have higher intraclass reliabilities for ten students. Moreover, the coefficient alpha and test-retest reliabilities for the scale scores were generally above .90 and .80 respectively. Student evaluation forms were administered over three semesters - spring and fall semesters of 1995, and spring semester of 1996. The sample of 21 institutions included similar numbers of two and four-year colleges and universities, although about half of the classes were from universities (largely comprehensive rather than research or doctoral universities). Table 1 indicates the breakdown by institutional type for this study. For analysis purposes, the classes were collapsed into the following eight discipline groups: health sciences (20 classes), business (75 classes), education (43 classes), social sciences (161 classes), fine arts (45 classes), natural sciences (167 classes), technology (32 classes), and humanities (173 classes). Female instructors taught 211 (28%) of the classes and male instructors taught 530 (72%) of the classes. Only in the health sciences were there more female than male instructors.

Table 1 - Institutional Type

Institutional Type	Number of Type	Percentage of Classes From Each Type
2 Year College	7	26%
4 Year College	8	25%
University	6	49%

Instrument

The student evaluation form used in this study was the Student Instructional Report II (SIR II), a new version of the original SIR which had been made available to colleges over the past 25 years by the Educational Testing Service (ETS). SIR II was developed to reflect more recent emphases in college teaching, such as active learning, course outcomes assessment, and the importance of student effort and involvement (Centra, 1998). This end-of-course survey consists of 45 questions grouped into eight scales: course organization and planning (5 questions); communication (5 questions); faculty/student interaction (5 questions); assignments, exams, and grading (6 questions); supplementary instructional methods (7 questions); course outcomes (5 questions); student effort and involvement (3 questions); and course difficulty, work load, and pace (3 questions). The development of these scales is described in the manual (Centra, 1998). An overall evaluation item and a student information section (5 questions) are also included in the form. With the exception of the course difficulty/workload/pace scale, the questions were evaluated using a Likert scale of 1-5. An "0" option was also available as a "nonapplicable" response. The response options for Scales A through E ranged from Very Effective to Ineffective. The response choices for Scales F and G ranged from Much More than most courses to Much Less than most courses. In a pilot test these options had proven more useful than an agree/disagree scale (Centra, 1998).

Data Analysis

Multivariate analysis of variance (MANOVA) was used to investigate the mean

differences on the several dependent variables simultaneously while controlling for the intercorrelation among them. By considering all the variables simultaneously, MANOVA becomes more powerful than performing separate ANOVAs on each of the variables.

Results

MANOVA results for the class mean scores on the eight SIR II scales and the overall evaluation item are presented in Tables 2 and 3.

Table 2 - MANOVA Results and Sample Sizes
Ratings of Female and Male College Instructors by Female and Male Students Within the Same Classes

	Sample Size		Instructor Gender	
	Female Instructor	Male Instructor	Female Instructor	Male Instructor
All Disciplines	N = 211	N = 530	2.68**	6.31***
Health	N = 14	N = 6	0.24	1.86
Business	N = 20	N = 55	1.99	1.25
Education	N = 9	N = 34	0.96	1.28
Social Sciences	N = 38	N = 123	1.27	2.44*
Fine Arts	N = 16	N = 29	1.38	1.81
Natural Sciences	N = 40	N = 127	1.89	2.14*
Technology	N = 8	N = 24	0.54	1.01
Humanities	N = 57	N = 116	1.00	1.64

Note: MANOVA results are represented by Wilks' Lambda
*p <= 0.05 **p <= 0.01 ***p <=0.001

Table 3 - MANOVA Results and Sample Sizes
Ratings by Female and Male Students of Instructors Across Different Classes

	Sample Size	Student Gender		Instructor Gender x Student Gender
		Female Student	Male Student	
All Disciplines	N =741	4.01***	4.79***	0.76
Health	N = 20	0.89	1.04	0.23
Business	N = 75	1.99*	1.65	0.69
Education	N = 43	1.14	1.26	0.27
Social Sciences	N = 161	1.61	2.55**	1.06
Fine Arts	N = 45	1.26	1.97	1.75
Natural Sciences	N = 167	3.85***	3.14**	0.76
Technology	N = 32	1.50	0.92	0.46
Humanities	N = 173	1.28	1.96*	0.57

Note: MANOVA results are represented by Wilks' Lambda
*p <= 0.05 **p <= 0.01 ***p <= 0.001

Separate MANOVAs were generated for female instructors, male instructors, female students, male students, and the interaction between instructor gender and student gender across each discipline and for all disciplines combined. Follow-up ANOVA tests were examined to determine on which scales instructors were rated differently based on student and instructor gender. Table 4 contains the mean ratings of

female and male instructors by female and male students. Asterisks identify significant ANOVAs for scales which indicate between instructor (gender) differences. Plus signs identify significant ANOVAs for scales which indicate between student (gender) differences. Results are indicated at the 0.05, 0.01, and 0.001 level of significance or better.

The first analysis examined the question, **Do female and male students in the same classes rate their instructors differently depending on the gender of the instructor?** MANOVA results are obtained from Table 2 and ANOVA results/mean scores are obtained by reading the vertical columns of Table 4 (plus signs indicate statistical significance). The mean scores being compared in this question are from female and male students within the same classes; thus the two groups of students are evaluating the same instructors.

Table 4 -Mean Ratings of Male and Female Instructors by Male and Female College Students (ANOVA Results)

Discipline		Scale A Course Organization & Planning MI		FI	Scale B Communication MI	FI	Scale C Faculty/Student Interaction MI		FI	Scale D Assignments, Exams, & Grading MI		FI	Scale E Supplementary Instructional Methods MI		FI	Scale F Course Outcomes MI	FI	Scale G Student Effort & Involvement MI	FI	Scale H Course Diff., Workload & Pace MI	FI	Overall Evaluation MI	FI
All Disciplines	MS	4.24	*	4.16 ++	4.27	4.21 ++	4.19		4.21 +	3.96		4.00 ++	3.91		3.92 ++	3.64	3.61	3.54	3.50 +	2.54	2.53	3.95	3.92 +
	FS	4.26		4.28	4.29	4.32	4.21	**	4.31	3.98	***	4.10	3.92	**	4.03	3.58	3.65	3.59	3.59	2.53	2.55	4.00	4.04
Health	MS	4.09		3.93	4.32	4.08	4.34		3.91	3.99		3.86	4.11		3.79	3.92	3.64	3.55	3.67	2.63	2.27	3.85	3.70
	FS	4.03		4.07	4.35	4.22	4.29		3.93	4.00		3.93	4.19		3.83	4.00	3.76	3.57	3.76	2.56	2.34	3.91	3.89
Business	MS	4.19		4.25	4.23	4.27	4.15		4.27	3.94		4.09	3.86	**	4.16	3.65	3.77	3.58	3.55	2.57	2.61	3.92	4.00
	FS	4.12		4.36	4.18	4.38	4.07		4.36	3.86	**	4.18	3.79	**	4.13	3.47	* 3.71	3.57	3.49	2.50	2.54	3.84	4.03
Education	MS	4.39		4.15	4.41	4.27	4.47		4.42	4.19		4.07	4.21		4.09	3.89	3.68	3.47	3.56	2.58	2.6	4.09	3.78
	FS	4.28		4.17	4.37	4.27	4.43		4.4	4.11		4.12	4.12		4.19	3.81	3.69	3.43	3.51	2.59	2.64	4.12	3.94
Social Science	MS	4.34	**	4.11	4.33	* 4.17	4.22		4.15	4.00		3.91	3.90		3.88	3.66	3.52	3.52	3.40	2.61	2.59	4.02	3.88
	FS	4.34		4.22	4.33	4.27	4.25		4.26	4.01		3.97	3.94		3.96	3.61	3.55	3.58	3.51	2.60	2.60	4.06	3.90
Fine Arts	MS	4.15 +		3.97	4.22 +	* 4.03 +	4.07 +		4.03	3.97		3.86	3.88	*	3.57 +	3.54	3.39	3.32	3.14	2.48	2.55	3.78 +	3.72
	FS	4.32		4.26	4.37	4.30	4.31		4.22	4.06		4.08	3.88		3.91	3.60	3.50	3.40	3.30	2.54	2.65	4.03	3.92
Natural Science	MS	4.17		4.2	4.19	4.22	4.10	*	4.28	3.85	*	4.02	3.81		3.85 +	3.49	3.57	3.65	3.6 +	2.39	2.4	3.88	4.02
	FS	4.21		4.32	4.22	4.33	4.12	***	4.4	3.88	***	4.17	3.82	***	4.06	3.44	* 3.64	3.73	3.82	2.39	2.41	3.91	* 4.13
Technology	MS	3.89		4.02	3.98	4.04	3.94		4.10	3.75		3.99	3.83		4.08	3.56	3.72	3.56	3.32	2.41	2.43	3.69	3.84
	FS	3.93		4.19	4.04	4.13	4.04		4.18	3.85	*	4.19	3.89	*	4.27	3.56	3.73	3.66	3.64	2.41	2.42	3.85	4.05
Humanities	MS	4.27		4.21	4.32	4.26	4.24		4.26 +	4.00		4.06	3.99		3.97	3.70	3.66	3.49	3.54	2.62	2.62	4.05	3.95 +
	FS	4.32		4.31	4.35	4.37	4.26		4.39	4.06		4.16	4.01		4.07	3.66	3.72	3.52	3.57	2.65	2.65	4.08	4.13

*p <= 0.05 **p <= 0.01 ***p<= 0.001 +p <= 0.05 ++p <= 0.01 +++p <= 0.001

* indicate significant ANOVAs for instructor gender differences + indicate significant ANOVAs for student gender differences

MI = male instructors FI = female instructors MS = male students FS = female students

Note: Scale H was collapsed into a scale of 1-3

BEST COPY AVAILABLE

The two MANOVAs for all disciplines combined indicated that there are significant differences in how female and male instructors were evaluated by female and male students. The MANOVA for female instructors (F = 2.68) was significant at the 0.01 level, while the MANOVA for the male instructors (F = 6.31) was significant at the 0.001 level (Table 2). The ANOVA results indicated that when female instructors were rated by female and male students, there was a significant difference in the mean ratings on six of the eight scales and the overall evaluation item, with the female students consistently awarding higher scores to female instructors. Mean scores for female instructors were: course organization and planning (FS M = 4.28; MS M = 4.16); communication (FS M = 4.32; MS M = 4.21); faculty/student interaction (FS M = 4.31; MS M = 4.21); assignments, exams, and grading (FS M = 4.10; MS M = 4.00); supplementary instructional methods (FS M = 4.03; MS M = 3.92); student effort and involvement (FS M = 3.59; MS M = 3.50); and overall rating (FS M = 4.04; MS M = 3.92). This indicates, of course, that female instructors received a lower rating from male students on these same scales. When male instructors were rated by female and male students, there were no significant ANOVAs, thereby indicating no significant difference in how the students of both genders evaluated the male instructors. In summary, female instructors received higher ratings on seven of the nine scales when rated by female students. Male instructors were not rated significantly different by female or male students.

The pattern of results at the various academic discipline level were more varied. The MANOVA for male instructors in the Natural Sciences (F = 2.14) was significant,

indicating that male instructors were rated significantly different by female and male students on the vector representing the eight scales and the overall evaluation item. There were, however, no significant ANOVAs. This finding is possible because the multivariate test considers the correlation among the variables and the joint differences on all the variables, while the univariate tests do not consider correlations among the variables and determines differences on each variable separately. The MANOVA for female instructors (F = 1.89) was not significant, thus indicating no differences between the ratings by female and male students. Under these circumstances, the ANOVAs should be interpreted cautiously since as the number of calculated F ratios increases so does the likelihood of obtaining a significant F by chance (two scales were significant at the .05 level). In summary, in the Natural Sciences there was generally little difference in how female and male students evaluated female and male instructors.

In the Social Sciences, there was a significant MANOVA for the ratings of male instructors by female and male students (F = 2.44), but not for the female instructors (F = 1.27). There were, however, no significant ANOVAs for the evaluation of male instructors by female and male students.

For the remaining six disciplines there were no significant MANOVAs for female or male instructors when rated by female and male students, although Fine Arts and Humanities reported some significant ANOVAs. In summary, female and male instructors did not differ significantly in their ratings by students in these six academic disciplines.

In the second analysis, the question examined was, **In general (i.e. across**

classes) do female and/or male students tend to give different ratings to female and male instructors?, MANOVA results are given in Table 3, and ANOVA results/mean scores are obtained by reading the horizontal rows of Table 4 (asterisks indicate statistical significance). The mean scores being compared in this question are from two groups of female students - one group evaluated female instructors and the other male instructors. Also, the mean scores for two groups of male students will be compared - one group evaluated female instructors and the other male instructors.

The two MANOVAs for all disciplines combined indicated that there were significant differences in how female and male students evaluated instructors. Both the female students' MANOVA (F = 4.01) and the male students' MANOVA (F = 4.79) were significant at the 0.001 level. The ANOVA results indicated that when female students evaluated female and male instructors, there was a significant difference in the mean ratings on three of the nine scales. In all three cases, the female students gave a higher rating to the female instructors than to the male instructors on faculty/student interaction (FI M = 4.31; MI M = 4.21); assignments, exams, and grading (FI M = 4.10; MI M = 3.98); and supplementary instructional methods (FI M = 4.03; MI M = 3.92). Male students differed significantly in their evaluation of female and male instructors on only one scale, giving male instructors higher ratings on course organization and planning (MI M = 4.24; FI M = 4.16).

In analyzing results from the various academic disciplines, the MANOVAs for the Natural Sciences were significant for both female (F = 3.85) and male students (F = 3.14).

Female students rated the group of female instructors significantly higher than the group of male instructors on four of the eight scales - faculty/student interaction (FI M = 4.40; MI M = 4.12); assignments, exams, and grading (FI M = 4.17; MI M = 3.88); supplementary instructional methods (FI M = 4.06; MI M = 3.82); and course outcomes (FI M = 3.64; MI M = 3.44) - and the overall evaluation item (FI M = 4.13; MI M = 3.91). Male students rated female instructors significantly higher than male instructors on two of the scales - faculty/student interaction (FI M = 4.28; MI M = 4.10); and assignments, exams, and grading (FI M = 4.02; MI M = 3.85). In summary, in the Natural Sciences female and male students frequently rated female instructors higher than male instructors. These results for female instructors are noteworthy in that Natural Science is a traditionally male dominated field. Additionally, in this study only 24% of the instructors in the Natural Sciences were women.

In the Social Sciences, the MANOVA for the evaluation of instructors of both genders by male students (F = 2.55) was significant, indicating an overall significant difference in how male students evaluated female instructors in comparison to male instructors. The ANOVAs indicated that male students rated the group of male instructors higher than the group of female instructors on two scales - course organization and planning (MI M = 4.34; FI M = 4.11); and communication (MI M = 4.33; FI M = 4.17). The MANOVA examining how female students (F = 1.61) evaluated female and male instructors was not significant. In summary for the Social Sciences, there were few significant differences between student genders; however, on two scales male students

rated male instructors higher than female instructors.

An analysis of the Business discipline indicated that only the MANOVA for female students (F = 1.99) evaluating the group of female and the group of male instructors was significant. The ANOVAs indicated that female students rated female instructors significantly higher than male instructors on three scales - assignments, exams, and grading (FI M = 4.18; MI M = 3.86); supplementary instructional methods (FI M = 4.13; MI M = 3.79); and course outcomes (FI M = 3.71; MI M = 3.47). In summary for Business, there were small differences but they were in favor of female students rating female instructors higher than male instructors on three scales.

Humanities reported one significant MANOVA indicating that male students (F = 1.96) evaluated female instructors, as a group, significantly different than the group of male instructors. There were, however, no significant ANOVAs.

The remaining four disciplines - Health (FS F = 0.89; MS F = 1.04), Education (FS F = 1.14; MS F = 1.26), Fine Arts (FS F = 1.26; MS F = 1.97), and Technology (FS F = 1.50; MS F = 0.92) - reported no significant MANOVAs for female students ratings of the group of female and the group of male instructors, or male students ratings of female and male instructors, although Fine Arts and Technology reported some significant ANOVAs. In these four disciplines, then, the ratings by female students of female and male instructors did not differ significantly. This is also true for the ratings by male students.

The lack of significant interactions in instructor/student gender for any of the disciplines individually or for all disciplines combined indicate largely main effect differences.

Discussion

The results of this study differ considerably from past studies which indicated either no differences or that male students rated male instructors higher than female instructors. Instead, female students in this study gave higher ratings to female instructors than to male instructor on three of eight scales for all disciplines combined; male students gave male instructors higher ratings on only one scale, course organization and planning. These results were based on the analyses across classes and there were no significant interactions. Within classes and for all disciplines combined, female students gave higher ratings than male students to female teachers on six scales and the overall evaluation item. But female and male students did not differ in their ratings of male teachers. For the total sample of classes, therefore, when more favorable ratings were given they were largely by female students to female instructors.

Is this gender bias in favor of female instructors or does it suggest that gender related teaching styles are affecting the evaluations, particularly since male instructors were not generally rated differently by either student gender. In Harris's (1975) study and in Feldman's (1993) meta-analysis teaching methodology rather than gender appeared to have a greater effect on evaluations. Belenky, Clinchy, Goldberger, and Tarule (1986) emphasized that female instructors, as well as female students, are more receptive to a

teaching methodology that values connection over separation, understanding and acceptance over assessment, and collaboration over debate. "Connected classrooms" provide an environment for growth and an acceptance of uncertainty since knowledge evolves over time and experience. Connected teachers, according to Belenky et. al., emphasize group work and discussions, and see their role as one of a facilitator. This methodology is in contrast to a more traditional, lecturing approach to teaching which Freire (1971) described as the "banking" method, in which the teacher's role is to "fill" the students by making deposits of information.

Supporting this interpretation are the higher evaluations that female instructors received on the faculty-student interaction and the supplementary instructional methods scales, in particular from female students who may respond more positively to "connected teaching."

It is interesting to note that in the Natural Sciences, which are largely male dominated (e.g. in this study three-fourths of the teachers were men), male students as well as female students rated women teachers higher on faculty-student interaction (also assignments, exams and grading). Similarly, in Business, another male dominated field (74% were men in this study), both male and female students rated women teachers higher on supplementary instructional methods. So in these largely male fields women are perceived by all students as having good rapport or using more active learning methods.

The stronger ratings that men received on course organization and planning,

particularly from male students in the social sciences, contributes to a theory of gender-related teaching strengths: Men are better organized in their courses but women are more connected to their students. Bennett's (1982) small sample study and Feldman's (1993) meta-analysis reached a similar conclusion.

Most of the statistically significant scale differences in this study were at least a half standard deviation or greater (s.d. ≅ 4.0 typically), indicating practical differences as well. But on the overall evaluation of instruction item, gender differences were generally smaller or not significant. This too supports the conclusion that gender differences among instructors was related more to their gender-related approaches to teaching than to their overall effectiveness.

References

Basow, S. A., & Distenfeld, M. S. (1985). Teacher expressiveness: More important for males than females? Journal of Educational Psychology, 77, 45-52.

Basow, S. A., & Howe, K. G. (1987). Evaluations of college professors: Effects of professors' sex-type, and sex, and students' sex. Psychological Reports, 60, 671-678.

Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? Journal of Educational Psychology, 79(3), 308-314.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). Women's ways of knowing. New York: Basic Books.

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. Journal of Educational Psychology, 74(2), 170-179.

Centra, J. A. (1998). Development of The Student Instructional Report II. Princeton, NJ: Educational Testing Service.

Elmore, P. B., & LaPointe, K. A. (1974). Effects of teacher sex and student sex on the evaluation of college instructors. Journal of Educational Psychology, 66(3), 386-389.

Etaugh, C., & Riley, S. (1983). Evaluating competence of women and men: Effects of marital and parental status and occupational sex-typing. Sex Roles, 9, 943-

952.

Feldman, K. A. (1993). College students' views of male and female college teachers: Part II-Evidence from students' evaluations of their classroom teachers. Research in Higher Education, 34(2), 151-211.

Freire, P. (1971). Pedagogy of the oppressed. New York: Seaview.

Harris, M. B. (1975). Sex role stereotypes and teacher evaluations. Journal of Educational Psychology, 67(6), 751-756.

Kaschak, E. (1978). Sex bias in student evaluations of college professors. Psychology of Women Quarterly, 2(3), 235-243.

Kaschak, E. (1981). Another look at sex bias in students' evaluations of professors: Do winners get the recognition that they have been given? Psychology of Women Quarterly, 5(5), 767-772.

Lombardo, J., & Tocci, M. (1979). Attribution of positive and negative characteristics of instructors as a function of attractiveness and sex of instructor and sex of subject. Perceptual and Motor Skills, 48, 491-494.

Paludi, M. A., & Bauer, W. D. (1983). Goldberg revisited: What's in an author's name. Sex Roles, 9, 387-390.